https://dx.doi.org/10.22161/ijaers.76.2 ISSN: 2349-6495(P) | 2456-1908(O)

Exploratory Study on the behavior of the Brazilian Financial Market using Google Trends

Fernando Gonçalves de Castro Filho, Alexandre Acácio de Andrade, Jadir Perpetuo dos Santos, Júlio Francisco Blumetti Facó , Sergio Goldman

Engineering, Modeling and Applied Social Sciences Center, Universidade Federal do ABC, Brazil

Abstract *— Search engines have changed the way people find relevant information on the Internet. It was suggested that investors use these Internet tools during the acquisition decision process and that the data sets generated by search engines are related to stock market movements. This study aims to analyze the relationship between this social phenomenon and the evolution of the Brazilian stock exchange. After previous studies, the correlation between Google's query volumes for terms related to a specific index of the Brazilian stock market (Ibovespa) and stock market metrics (opening, closing, high and low prices) was analyzed. The data showed a positive correlation between investor attention, represented by Google's search volumes and market prices. These results suggest that Google Trends data is more positively correlated to the highs of financial data, suggesting that investors tend to search more on the web when the market reaches its peak or vice versa. This article highlights the potential that this source of information has in understanding the Brazilian stock market.*

Keywords *— **Brazilian Financial Market, Data analysis, Google Trends.***

I. INTRODUCTION

Financial market prices fluctuate at all times, and this is the result of interaction between participants in this market. The participants can range from companies with large capital to small individual investors, but they all buy or sell financial products according to their expectations about the future of the market.

Usually, what is expected of the consumer - and, in particular for this study, we can substitute the term consumer for investor - is that he always seeks to increase his satisfaction by reducing any uncertainty involved in the process of choosing a product, or any investment. Thus, consumers and investors tend to assume a more active behavior in the search for information related to that product or investment, which, nowadays, thanks to technological advances and greater accessibility, leads to searches on the Internet. When a consumer searches for products online, this usually leads to the purchase of this product [35].

According to Simon (1955), the investor's decision-making process begins with the stage of gathering information. With the use of the Internet, the process of collecting information has been greatly facilitated over time, and search engines are the main technology for this purpose. This type of technology helps in web browsing and, many users use it as an entry point to the World Wide Web [24].

As of 2006, Google (considered one of the largest search engines in the world) allowed any interested party to access information regarding the frequency of specific search terms that were being inserted into its search engine. An innovative tool - called Google Trends - its data originates from the constant use of the Google tool -, has been used as a way to explain some social phenomena; a simple example is the detection of influenza epidemics according to the search for words related to the symptoms of the disease [17]. This type of data has also been applied to the most varied subjects, from forecasts related to the unemployment level in the United States [12] to forecasts of house and apartment prices [39].

These types of analysis have been used to forecast gold prices [26], to forecast commodity movements [34] and to forecast general market volatility [10]. According to Da & Gao (2011), the volume of searches carried out on the Internet related to stock market index names can be interpreted as a measure of investor attention to this market. When this investor finds increasing demand for information about a given market index, he has a great chance of using search engines as a source of information.

https://dx.doi.org/10.22161/ijaers.76.2 ISSN: 2349-6495(P) | 2456-1908(O)

Thus, analyzes that establish a relation between databases of search volumes and the behavior of the financial market in Brazil have been little explored. This study aims to fill part of this gap and provide more depth to studies aimed at this market, drawing parallels between studies consolidated in the United States and bringing them to the reality of the Brazilian stock market.

Given this scenario, the present study seeks to understand whether it is possible to use the search data, generated on the web, as a source of information to predict the behavior of the Brazilian financial market. Following the recent literature ([31]; [7]; [33]), it can be said that this type of data can be used for the financial market in the United States, however, what we propose here is to explore these same properties in a smaller market, in this case, Brazil; which remains underdeveloped and is more dependent on foreign capital. the paper.

II. LITERATURE REVIEW

This item represents the literature review, considering themes that helped to understand the behavior of investors in these markets, explaining the possible reasons for the adoption of certain actions and reactions.

Considering the Brazilian market as a whole, there are few studies that relate the behavior of Brazilian consumers with the search volumes obtained by search engines, this additional reading was carried out on the topic, as shown in Table 1.

Table.1: Works carried out relating the search volume in search engines with financial instruments.

Title	Author	Year	Search mechanism	Market	Financial Instrument	Statistical Tool
The Forecasting Power of Internet Search Queries in the Brazilian Financial Market	Henrique Pinto Ramos, Kadja Katherine Mendes Ribeiro e Marcelo Scherer Perlin	2017	Google	Brazil	PETR4, VALE5, BVSP, Taxa CDI, Selic, Tesouro Direto and Renda Fixa	Auto-regressive vector model (VAR) and Granger causality tests
Can internet search queries help to predict stock market volatility?	Thomas Dimpfl e Stephan Jank	2016	Google	USA and Europe	DJIA, FTSE 100, CAC 40 and DAX	Volatility models with and without research queries, Granger causality test
Quantifying the semantics of search behavior before stock market moves.	Chester Curme, Tobias Preis, H. Eugene Stanley e Helen Susannah Moatb	2014	Google	USA	S&P 500	Own modeling - relative change in search volume - compared to benchmarks and calculating
Quantifying trading behavior in financial markets using Google Trends.	Tobias Preis, Helen Susannah Moat, e H. Eugene Stanley.	2013	Google	USA	DJIA	Own modeling - relative change in search volume - compared to benchmarks
Modeling movements in oil, gold, forex and market indices using search volume index and twitter sentiments.	Tushar Rao e Saket Srivastava	2013	Google	USA, Commodities and Forex	DJIA, NASDAQ-100, Oil, Gold and EUR/USD (Forex)	Pearson's correlation, cross-correlations and Granger's causality test

https://dx.doi.org/10.22161/ijaers.76.2 *ISSN: 2349-6495(P) | 2456-1908(O)*

Web Search Queries Can Predict Stock Market Volumes.	Ilaria Bordino, Stefano Battiston, Guido Caldarelli, Matthieu Cristelli, Antti Ukkonen e Ingmar Weber	2012	Yahoo	USA	Listed companies - NASDAQ 100	Correlation between query volume and volatility, cross-correlation, Granger causality test, resampling
Predicting financial markets: Comparing survey, news, twitter and search engine data.	Huina Mao, Scott Counts e Johan Bollen	2011	Google	USA and Commodities	DJIA, VIX and Gold	Pearson correlation, Granger causality test and multiple linear regression
In search of attention.	Huina Mao, Scott Counts e Johan Bollen	2011	Google	USA	Russell 3000 and IPO	Pearson's correlation, autoregressive vector model (VAR)
Complex dynamics of our economic life on different scales: insights from search engine query data.	Tobias Preis, Daniel Reith e H. Eugene Stanley	2010	Google	USA	S&P 500	Cross-correlations and autocorrelations

Source: The Authors, 2020

2.1 Search engines

Search engines are tools designed to retrieve information (Information Retrieval) from the World Wide Web. The input of a search engine is usually a query that consists of one or more keywords creating a query term. The output can be web pages, images, or other types of files. These responses are presented on a search engine results page (SERP) and are called hits. The SERP is a list of hyperlinks that point to the pages found along with descriptive text (snippet or teaser). The database used to store information on the web page is generally called an index [24].

The motivation for using search engines comes from the sum result of several factors. The search process begins with the need for information. The user is looking for information and, from there, formulates the query in its verbal form, choosing some keywords. The system then returns to those files that, within a collection of documents, correspond to the query [24].

Brin & Page (1998) mention that people, even at the beginning of the 21st century, were used to surfing the web using links, usually starting with hierarchical indexes maintained by companies - such as the Yahoo portal. This was about to change when link analysis engines were introduced to the world of Information Retrieval. This technique consists of exploring the additional information that the web's hyperlink structure has to improve the search results [23].

To better contextualize the impact that the link analysis mechanisms have brought to this market, the Google search engine - founded in 1998 by the authors of the previously cited article, Sergey Brin and Lawrence Page -, were the pioneers in the use of this type of analysis. In mid-2004, this same company already had the largest market share (37%) among search engines at the time, followed by the Yahoo conglomerate (27%). In 2020, Google dominates the search engine market representing 72% of all searches made on the web, followed by the Chinese search engine Baidu with 14%, as illustrated in Table 2 [28], etc.

Table 2: The 10 largest virtual search engines[28]

Positon	**Search Engine**	**Share %**
1	Google	91.17
2	Bing	3.12
3	Yahoo!	3.06
4	Baidu	0.77
5	YANDEX RU	0.36
6	Ask Jeeves	0.26
7	YANDEX	0.23
8	Naver	0.14
9	AOL	0.13
10	Haosou	0.1

https://dx.doi.org/10.22161/ijaers.76.2 *ISSN: 2349-6495(P) | 2456-1908(O)*

In 2018, Google processed an average of 6 billion searches per day and 2 trillion searches per year worldwide. This corresponded to approximately more than 70,000 search queries each second. Another interesting fact is that in August 2012, the company's former senior vice president and responsible for the development of its search engine revealed that Google found more than 30 trillion unique URLs on the Web, crawled 20 billion sites per day and processed 100 billion searches per month [38].

After expanding significantly in the first decade of the 21st century, the growth rate of Google's search volume began to decline in 2009 and 2010, currently estimated at around 10% per year. In the start-up phase, growth was phenomenal, with an increase of 17,000% per year. The volume of research in the period 1998 - 1999 reached 1000%; in the 1999 - 2000 range it reached 200%. Google searches continued to grow at rates between 40% to 60% from 2001 to 2009, and that's when it started to slow down, stabilizing at a rate of 10% to 15% in recent years[22].

2.2 How search engines work

According to Lieberam-Schmidt (2010), search engines can be classified in two ways: between portals and result providers or between web catalogs or search engines. In the first classification set, queries are evaluated by how they are processed; for the second set, how the databases are generated are measured. It is worth mentioning that many of the existing research tools cannot be classified into just one of the categories. They use, for example, directory search results and search engine on the same results page.

The weblink structure is used as an evaluation criterion and the academic citation process comes in as inspiration to evaluate the importance of the pages. What the algorithm does is to count "citations" or links that a page has, both for reference to other sites and other sites pointing to the page. The rationale is: the greater the number of references that a given page presents, the greater the chance that this page will be relevant and, consequently, it will be presented to the user first.

In the article in which the Google tool originated, Brin & Page (1998)[4] explain some differences and improvements that the search engine presents concerning the “conventional” structure (Figure 1). One of them is the PageRank, which for the present study, will be classified as an analyzer.

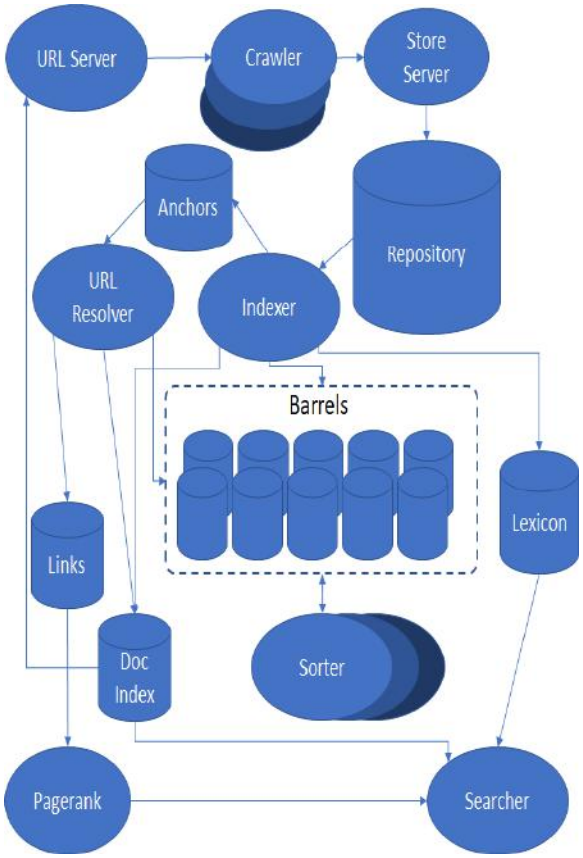

Fig. 1: High-level Google architecture [4]

2.3 Estimation and search data

Whenever a user accesses a search engine and makes a query, data such as the day, time, search term, and the region where the computer is located are saved in databases of the companies that provide these services. These data can be classified as search data on the web and are called digital trace.

These " digital trace " that are left can provide insights into a wide range of disciplines [3] and can apply to virtually any market where Internet searches precede the transaction, which holds for a large share of the world economy [39].

According to Oust & Martin (2018), inserting Google search data into prediction models for the real estate market improves its performance. The author explains that the choice of search terms is a fundamental part of the success of the model, as well as the choice of the region where the terms are searched.

According to Limnios (2018), increasing models using data from Google Trends does not improve its performance in predicting movements in the real estate market, and the use of this practice worsens the model's forecasting capacity. The same result was shown by Oust & Martin

https://dx.doi.org/10.22161/ijaers.76.2 *ISSN: 2349-6495(P) | 2456-1908(O)*

(2018), when he replaced an established market indicator - the Consumer Confidence Index (CCI) started in 1977 - with the volume of Google searches, with a worsening of the model's assertiveness.

2.4 Financial market

At all times, economic agents are making decisions about how they will consume, as well as their production of goods in the capital market. Within this group of agents, there will always be those who consume more than they produce, and those who produce more than they consume. In this way, the second group (savers) can meet the needs of the first group (borrowers) with what was saved. Therefore, the creation of a mechanism that makes the transfer of these resources feasible has become fundamental for the development of productive activities in modern society [15].

2.5 Capital market - Brazil

In Brazil, all operations that take place in the financial market and their respective participants are supervised by the Securities and Exchange Commission (CVM). The CVM was created by Law No. 6,385, on December 7, 1976, to discipline, supervise, and develop the securities market (https://www.investidor.gov.br/menu/Menu_Investidor/introducao_geral/introducao_mercado .html). It is an autarchy under a special regime directly linked to the Ministry of Finance and its scope can be seen in Figure 2[8].

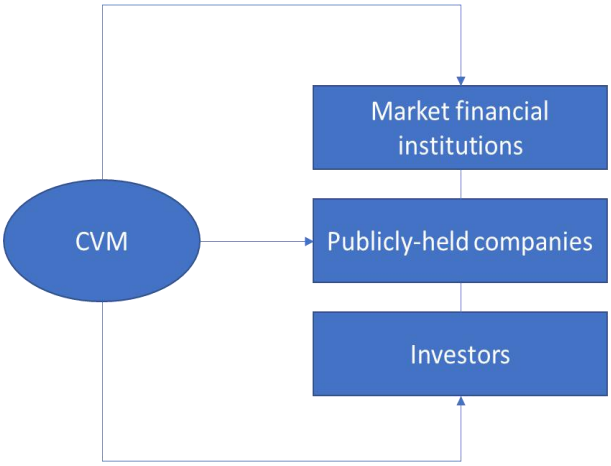

Fig. 2: CVM's scope [15]

The largest stock exchange in Latin America, B3 (former Bovespa), is located in Brazil. It has its headquarters in the central region of São Paulo, also the largest financial center in the region. This is the only stock exchange in the country for trading securities and it has listed 267 companies (in April 2020). Shares in the Brazilian market, different from what is commonly used in other international markets, can take two formats: common shares or preferred shares.

Usually, when one wants to evaluate the performance of investments involving shares, the Bovespa index always indicated as a reference for such comparison, being considered the main benchmark for the variable income market. The Ibovespa (Bovespa index) is revalued every four months and is the result of a theoretical portfolio of assets. Its main objective is to be the indicator of the average performance of the most tradable and representative assets of the Brazilian stock market (approximately 80% of the number of trades and the financial volume of the entire market), being composed only by the shares and units of companies that meet the criteria of its methodology [1].

Another interesting feature of the Brazilian capital market is the distribution of the types of investors and their respective market shares. It can be seen from Figure 3; currently, half of the Brazilian spot market is composed of foreign investors. These showed strong growth between the years 2010 to 2014, and have maintained steady growth since then. Only 17.2% of the market share is held by individual investors [2].

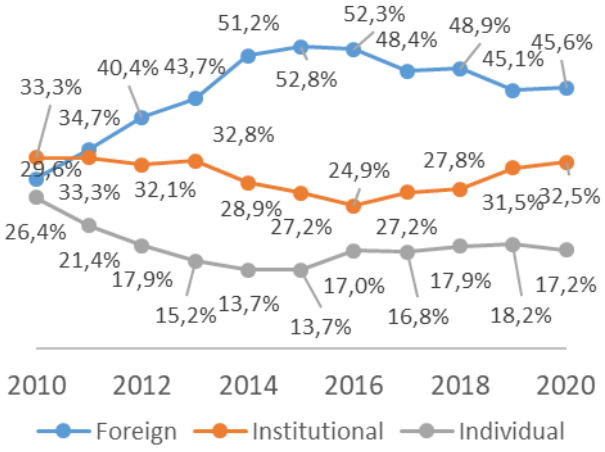

Fig. 3: Evolution of Investor Participation in the Market from a Brazilian perspective.[2].

III. METHODOLOGY

The present study has an exploratory character, and it aims to provide an overview of a subject, until then, little-explored to our national market. Descriptive, as it presents characteristics of a certain phenomenon and establishes a relationship between its variables. It is slightly explanatory, as it is also concerned with identifying factors that contribute to the occurrence of the studied phenomenon [16].

https://dx.doi.org/10.22161/ijaers.76.2 *ISSN: 2349-6495(P) | 2456-1908(O)*

The scientific methodology used is the result of the combination of the inductive method and the statistical method. The first method will provide the logical basis for the investigation being proposed, starting from particular cases to generalization [13]. The second method, based on the application of variable statistical theories of the phenomenon[41], making it possible to determine, in numbers, the probability of a given conclusion[16].

According to Popper (2015), the inductive leap from particular cases to generalization requires that observations of particular phenomena reach infinite proportions, which never occurs. For Gil (2019), one way to get around this problem is to use the statistical method, which makes it possible to indicate different degrees of confidence for a conclusion obtained through induction. The limitation of the statistical method is because its explanations cannot be considered as absolute truths, but as explanations with good chances of being true.

According to Dimpfl (2016), who also assumes that search engines are the primary source of information for amateur investors, the authors demonstrate that an increase in the number of searches performed on search engines today is followed by increased volatility on the markets tomorrow.

To corroborate the aforementioned hypothesis, the web search data generated by the Google search engine tool will be used to estimate the movements of the Brazilian stock market, concerning asset pricing. Google will be used due to its dominance in the search engine market (See Table 2 - The 10 largest virtual search engines), thus being the best available representation of the general behavior of the population in this type of tool.

Thus, to apply the methodology for assessing the validity of this hypothesis and its causal relationships, the following steps presented in figure 4 will be followed.

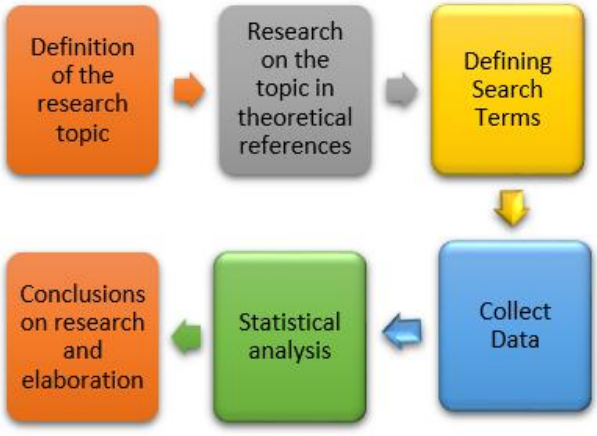

Fig. 4: Methodological Flow

IV. DEVELOPMENT OF THE RESEARCH

Next, the steps of the research and its data analysis shall be presented.

4.1 Definition of search terms

The search terms are the keywords that the user enters into the search engines. The whole process begins at the moment when the user identifies a research task to be accomplished and converts this need into words. In order to achieve success and obtain relevant information, it is essential to choose the correct search term, otherwise, the results obtained may be unsatisfactory [24].

We can illustrate this process using a simple example: a potential investor in the company "X" wants to obtain information about the value of its shares. This way, he inserts "X" as his keyword in the search engine and, as a result, obtains the link to the company's website, its geographical location and its social networks. For this specific task (obtaining the share price of company "X"), the key term could be the ticker itself (an abbreviation used to exclusively identify shares traded on the stock exchange, eg "ITUB3" for Banco Itaú and "ABEV3" for the manufacturer AmBev beer) of company "X" or even the addition of the term "shares" before the company name.

Lieberam-Schmidt (2010) explains that during the research process, some common mistakes lead to unsatisfactory search results. They are: I. believing that information is necessary for a task and this is not true; II. verbalization of the task does not reflect the need for information; III. poor formulation of the consultation; IV. the same word has different meanings (polysemy) and V. use of synonyms.

For the more specific analysis of this market, only six (6) companies were selected to facilitate the methodology, data analysis[14], and the amount of information. For this selection, the companies with the greatest representation in the market were listed, as well as the differentiation of the sectors, which will be specified according to each company analyzed. The final list of search terms for individual companies as shown in Table 3.

https://dx.doi.org/10.22161/ijaers.76.2 *ISSN: 2349-6495(P) | 2456-1908(O)*

Table 3: List of key terms used for individual companies separated by economic sector and ordered by participation in IBOVESPA [20].

#	Company Name	Sector	Part. (%)
1	Vale	Mining	12,89
2	Itaú	Financial Intermediary	10,89
3	Petrobras	Oil, Gas, and Biofuels	7,48
4	Renner	Retail	1,75
5	Cogna	Education	1,28
6	Gerdau	Metallurgy	1,41
Representativeness on the IBOVESPA			35,33

The selection made represents 6 different sectors of the Brazilian economy and approximately 35% of the Ibovespa. As previously stated, for the Bovespa index, the term "Ibovespa" will be used.

Given all the data collected so far, and to have a faster and more objective targeting of them, tickers were not used in searches on Google Trends, but rather the names of the companies themselves.

4.2 Data collection

Two types of data were used during the study, these are search data and financial data. The period evaluated will be from January 2008 to January 2018. The features and particularities of the tooling will also be described, when necessary.

4.2.1 Search Data

According to its question page, the Google Trends service analyzes a random part of searches performed on Google itself to calculate how many searches were made for specific terms. The tool is based on two variables: time and location. Each data point in Google Trends is calculated by dividing the total searches for a specific geographic region and the length of time the total searches cover, which results in a measure of relative popularity. The resulting numbers are scaled in a range from 0 to 100 based on the ratio of one topic to all searches across all topics. For example, a value of 100 represents the peak popularity of a term[18].

It's important to keep in mind that Google Trends only provides data related to search terms for which traffic exceeds a certain threshold, disregarding those with low volume. Repeated queries by a single user for a short period of time are also excluded so that the level of interest is not artificially impacted by this type of behavior. There is also the query filter with apostrophes or other special characters.

Google Trends does not provide search volume with daily granularity, except for extremely frequent search terms, this is not the case for all of our terms[18]. Therefore, we conduct our analyzes on a weekly granularity, thus, data related to the set of selected search terms are available. Research data was also obtained using region filters, both for Brazil and for Global. What we want to validate here is whether there is any difference between the research relationships in the different regions, given that the Brazilian financial market has a strong influence from foreign investors, another important limitation to be highlighted is that the Google Trends values vary according to the date of access to the tool, due to its normalization process. For example, when a new top of interest is reached, it becomes the new value "100" in the database, changing everything else. In addition to this, the normalization process results in only whole numbers, which can also add rounding errors. These problems do not currently have a solution and can impair the reproducibility of the results [6].

In the case of Brazil, the dominance of the Google search engine is clear. In the country, for the year 2020, the mechanism had 97.33% of the national search market, against timid 1.33% of the second-place Bing.[37].

4.2.2 Financial data

In this work, weekly trading volumes for the general Ibovespa market index and weekly volumes for the first 3 individual companies listed in figure 6 were used as inputs.

The volumes traded in isolation were analyzed, the direction (ascending or descending) that the market presented during that determined period was disregarded. This evaluation aimed to verify if there is any type of correlation between the volume of searches and the volume of negotiations that occur for a financial instrument - which aims to be a reference for investors – [7] and, for individual companies [33].

Weekly opening and closing prices were also used for both the Ibovespa and the 6 (six) largest individual companies listed on this list. The other possibilities would be to use the weekly maximum or minimum price, however, we opted for the analysis of opening and closing prices by what they represent for the market.

According to Elder (2004), the opening price generally reflects the opinion of amateur investors about the value. Opening prices, most of the time, are closer to the lows or

https://dx.doi.org/10.22161/ijaers.76.2 *ISSN: 2349-6495(P) | 2456-1908(O)*

highs of that period, and during the course of the day, the asset price tends to recover and reach a more moderate value.

Still, according to Elder (2014), the closing price tends to reflect the opinions of professional investors, characterizing them as more active at the end of the trading periods. For the author, professional investors generally operate against amateur investors. The first group tends to profit according to the more emotional reactions of the second group.

All financial information (volume, opening, and closing price) presented in figure 5, were extracted from the Yahoo Finance portal. The service is provided free of charge and, in this way, the website informs that the data provided is for informational purposes only and is not intended for commercial or investment purposes [40]. Another feature of the service is that it receives information from the data company ICE Data Services (market data company) and passes it on in the original format received with a 15-minute delay (for Brazil).

Itaú Unibanco Holding S.A. (ITUB4.SA) ☆ Adicionar à lista
Sao Paolo - Sao Paolo Preço Adiado. Moeda em BRL.

47,93 -0,05 (-0,10%)
No fechamento: 19 de outubro 5:09PM BRT

Resumo | Gráfico | Conversas | Estatísticas | Dados Históricos | Perfil | Finanças | Análises | Opções | Titulares | Sustentabilidade

Período: 01 de jan de 2008 - 01 de jan (| Mostrar: Preços históricos | Frequência: Semanalmente | Aplicar

Câmbio em BRL | Fazer download dos dados

Data	Abrir	Alto	Baixo	Fechamento*	Fechamento ajustado**	Volume
25 de dez de 2017	42,54	42,99	42,39	42,58	42,21	19.160.100
18 de dez de 2017	41,00	42,83	40,67	42,53	42,16	45.240.100
15 de dez de 2017				**0.1445** Dividendo		
11 de dez de 2017	42,21	43,10	40,77	41,08	40,58	115.043.200
04 de dez de 2017	41,84	42,78	40,77	42,36	41,85	46.377.900
01 de dez de 2017				**0.015** Dividendo		
27 de nov de 2017	42,99	43,33	40,70	41,84	41,32	53.264.600
20 de nov de 2017	42,98	43,80	41,95	42,69	42,16	32.499.600

Fig. 5: Example search result, using Yahoo Finance, for Banco Itaú (ITUB4) for the trial period

In this research, the programming language interpreted by Python was used to carry out statistical analysis. Python is a free, open-source language, and an excellent option as the main language for creating data applications [27].

The causality study seeks to identify whether a given variable is capable of improving the prediction of another in the context of time series. In order to ascertain whether this causality relationship, Granger's concept of causality (1969) was used.

Granger causality direction can be defined as:

a) Unidirectional (PF ⇒ PV);

b) Bilateral Causality (PF ⇔ PV);

c) Independence (PF and PV are independent, PF not Granger causes PV and PV not Granger causes PF).

According to Gujarati (1999), if the calculated probability value exceeds the critical value at a chosen level of significance, we reject the null hypothesis. That is, if the null hypothesis is rejected, it can be said that there is a causal relationship of Granger between the variables.

All of these statistical analyzes will be presented according to the companies mentioned and developed. Below, Granger's concepts, as well as proportionalities,

https://dx.doi.org/10.22161/ijaers.76.2 *ISSN: 2349-6495(P) | 2456-1908(O)*

will be exemplified, commented, and analyzed according to the assets situation of each institution in the defined period from 2008 to 2018.

4.2.3.1 Data from Vale

Vale, a Brazilian multinational mining company and one of the largest logistics operators in the country. It is one of the largest mining companies in the world and also the largest producer of iron ore, pellets, and nickel. The company also produces manganese, ferroalloy, copper, bauxite, potassium, kaolin, alumina, and aluminum. On the Stock Exchange (see Table 4) it remains listed on B3 with common shares (VALE3) and is also present in the fractional market (VALE3F).

Table 4: Descriptive analysis of Vale's asset prices.

VALE				
	Opening	**Maximum**	**Minimum**	**Closing**
Min.	9.63	13.03	8.60	**9.72**
Median	36.45	38.96	33.54	**36.49**
Max.	65.54	73.80	64.65	**65.64**

4.2.3.1.1 Vale statistical analysis

According to Table 5, all correlations of Vale's assets remain positive and of low strength, and there is a remarkable correlation of greater expression in the maximum price of the asset when compared to data from Google Trends do Brazil and a correlation of greater expression in the opening price (Open) of the asset when compared with data from Google Trends of the World, indicating that the interest in searching for the asset when it reaches its maximum value in the current month for the Brazilian market and the opening for the Global market.

Table 5: Vale correlations.

	CORRELATIONS WITH ASSETS VALE			
	Opening	**Maximum**	**Minimum**	**Closing**
WORLD	0.4959763	0.4941035	0.4510825	0.462266
BRAZIL	0.5349776	0.5398466	0.4782662	0.502440

According to the test for trends without a geographic filter (world) it is possible to establish a causal Granger relationship between prices (opening, closing, maximum and minimum) and search interest in trends with up to 99% confidence. This relationship is not possible to prove when significance levels below 0.1% are adopted except for the opening price (Open).

With no acceptable significance level, a causal relationship of trends with prices was found (Table 6), that is, trends do not precede price.

Table 6: Granger Causality - World Vale.

VALE - WORLD			
GRANGER CAUSALITY TEST - Probabilities			
	Price -> Trends		**Trends -> Price**
Opening	0.0006915095	↔	0.2597236
Maximum	0.001761687	↔	0.8140406
Minimum	0.001761687	↔	0.7988959
Closing	0.0011733033	↔	0.7370794

According to the test for trends without a geographic filter (world) it is possible to establish a causal Granger relationship between prices (opening, closing, maximum and minimum) and search interest in trends with up to 99% confidence. The present relationship is not possible to prove when significance levels below 0.1% are adopted.

With no acceptable significance level, a causal relationship between trends and prices were found (Table 7), that is, trends do not precede price.

Table 7: Granger Causality - Brazil Vale.

VALE - BRAZIL			
GRANGER CAUSALITY TEST - Probabilities			
	Price -> Trends		**Trends -> Price**
Opening	0.001125861	↔	0.2370143
Maximum	0.001942837	↔	0.5780904
Minimum	0.00327824	↔	0.2700823
Closing	0.002488181	↔	0.3014495

4.2.3.1.2 Statistical analysis ITAÚ

Itaú Unibanco Holding S.A. In 2019, Itaú Unibanco had R $ 28.3 billion in net income and distributed R $ 18.8 billion in dividends.

https://dx.doi.org/10.22161/ijaers.76.2 *ISSN: 2349-6495(P) | 2456-1908(O)*

Itaú Unibanco is a major merger of companies with numerous employees. Due to the union between Casa Moreira Salles, which in 1960 became Unibanco; and Banco Central de Crédito S.A., which was renamed Itaú.

The two banks merged in 2008. This merger led to the Instituto Moreira Salles, Instituto Itaú Cultural, Fundação Itaú Social, and Instituto Unibanco. With around 5,000 branches in Brazil and abroad, it also operates in 21 countries.

Itaú shares (ITUB4, ITUB3) are part of the IBOVESPA index and were one of the sensations from 2016 to 2019, going from the range of R $ 20 then to the current R $ 35/36 now, an appreciation of almost 100% in three years. An investor who had purchased R $ 100 thousand in Itaú shares (ITUB4, ITUB3) in 2016 would have earned R $ 75 thousand by December 2019.

In table 8 we can see that for the analyzed period, the median closing value of Banco Itaú shares is closer to its minimum value than to its maximum value, this can show that the prices of this share, in moments of euphoria, tend to value themselves more strongly than to devalue themselves in times of pessimism.

Table 8: Descriptive analysis of Itaú asset prices.

COMPANY: ITAÚ				
	Opening	**Maximum**	**Minimum**	**Closing**
Min.	8.94	10.94	7.00	**9.19**
Median	17.73	17.87	15.75	**16.97**
Max.	36.80	43.82	35.35	**43.74**

4.2.3.2.1 Itaú statistical analysis

According to Table 9, all correlations of the Itaú asset remain negative and of very low strength, thus indicating that the interest in searching for the asset is independent of prices and vice versa.

Table 9: Descriptive analysis of Itaú asset prices.

	CORRELATIONS WITH ITAÚ ASSETS			
	Opening	**Maximum**	**Minimum**	**Closing**
WORLD	-0,253752	-0.250883	-0.228290	-0.234694
BRAZIL	-0.359435	-.3590946	-0.332986	-0.341298

Unlike all the assets studied so far, there is no causal relationship of Granger (Table 10), either from price trends or from prices to trends at any level of acceptable significance.

Table 10: Granger Causality - Itaú Mundial.

ITAÚ - WORLD			
GRANGER CAUSALITY TEST - Probabilities			
	Price -> Trends		**Trends -> Price**
Opening	0.2163918	↔	0.7464781
Maximum	0.1119121	↔	0.6729117
Minimum	0.39484	↔	0.4449923
Closing	0.1541122	↔	0.4560661

Even in the face of the assets World against Brazil, Granger's theory of non-causality for both is confirmed - being the first (and only) company to behave in this way. Table 11 confirms the theory.

Table 11: Granger Causality - Itaú Mundial.

ITAÚ - BRAZIL			
GRANGER CAUSALITY TEST - Probabilities			
	Price -> Trends		**Trends -> Price**
Opening	0.2488691	↔	0.4629911
Maximum	0.2166089	↔	0.3747098
Minimum	0.3994947	↔	0.5587284
Closing	0.2385429	↔	0.2839803

4.2.3.2.2 PETROBRÁS statistical analysis

For four decades, since its creation in 1953, Petrobras has monopolized oil research, refining, and transportation in Brazil. In 1997, it lost this condition when Law 9.478 allowed other companies based in Brazil to start operating in all stages of the oil chain.

Even so, Petrobras still maintains majority control in the fuel production chain, and is one of the most important companies in the energy sector in the world, specializing in oil production in deep and ultra-deep waters.

Shaken by the corruption scandals pointed out by Operation Lava Jato, (Infomoney, 2020) the oil giant has accumulated losses of more than R $ 70 billion in successive losses since 2013, interrupted in 2018. Officially, it recognizes the loss of around R $ 6 billion for

https://dx.doi.org/10.22161/ijaers.76.2 *ISSN: 2349-6495(P) | 2456-1908(O)*

corruption between 2004 and 2012 and ensures that its new governance model reverses the problems highlighted by Lava-Jato operation.

The publicly traded mixed-capital company is controlled by the Federal. It has more than 60 thousand employees and produces around 1.7 thousand barrels of oil products daily. Its capital consists of common (PETR3, PBR, and ADR) and preferred (PETR4, PBR / A-ADR) shares traded on the São Paulo, New York, Madrid, and Buenos Aires stock exchanges.

In table 12 we can observe a phenomenon similar to what occurred in the descriptive analysis of Itaú's actions. During the analyzed period, the median closing value of Petrobras shares is closer to its minimum value than to its maximum value, this may show that the prices of this share tend to appreciate more strongly in times of euphoria than to devalue in moments of economic pessimism.

Table 12: Descriptive analysis of Petrobrás asset prices.

COMPANY: PETROBRAS				
	Opening	**Maximum**	**Minimum**	**Closing**
Min.	4.74	5.28	4.12	**4.84**
Median	19.76	21.58	17.90	**19.72**
Max.	50.97	53.68	43.00	**49.00**

According to Table 13, all correlations of the Petrobras asset remain positive and of moderate strength and there is a notable correlation of greater expression in the maximum price of the asset when compared to data from Google Trends do Brazil. Indicating that the search interest for the asset when it reaches its maximum value in the current month.

Table 13: Petrobras Correlations

	CORRELATIONS WITH PETROBRAS ASSETS			
	Opening	**Maximum**	**Minimum**	**Closing**
WORLD	0.6003359	0.6111687	0.5811080	0.5883111
BRAZIL	0.6573648	0.6725125	0.6383784	0.6546579

According to the test for trends without a geographic filter (world) it is possible to establish a causal Granger relationship between prices (opening, closing, maximum and minimum) and search interest in trends with up to 95% confidence. The present list is not verifiable when significance levels below 0.1% are adopted except for the opening price (Open).

With no acceptable level of significance, a causal relationship (Table 14) of trends with prices was found, that is, trends do not precede price.

Table 14: Granger Causality - Petrobrás Mundial

PETROBRAS - WORLD			
GRANGER CAUSALITY TEST - Probabilities			
	Price -> Trends		**Trends -> Price**
Opening	0.0009736962	↔	0.1506712
Maximum	0.004103743	↔	0.3344646
Minimum	0.006109518	↔	0.2525364
Closing	0.01791153	↔	0.3277053

According to the test for trends without a geographic filter (world) it is possible to establish a causal Granger relationship between prices (opening, closing, maximum and minimum) and search interest in trends with up to 95% confidence. The referring relation is not subject to proof when significance levels are adopted below 1% except for the opening (Open) and maximum (High) price.

With no acceptable significance level, a causal relationship (Table 15) of trends with prices was found, that is, trends do not precede price.

Table 15: Granger Causality - Petrobrás Brazil

PETROBRAS - BRAZIL			
GRANGER CAUSALITY TEST - Probabilities			
	Price -> Trends		**Trends -> Price**
Opening	0.001460834	↔	0.06837931
Maximum	0.006888635	↔	0.1675851
Minimum	0.01052826	↔	0.2512003
Closing	0.02282375	↔	0.5383188

https://dx.doi.org/10.22161/ijaers.76.2 ISSN: 2349-6495(P) | 2456-1908(O)

V. CONCLUSION

After all this analytical historical survey, researched from the purpose of verification when using query data for searches generated by search engines as a source of information to predict the behavior of the Brazilian financial market, it was possible to understand the existence of a certain consistency in the data which point to confirming Granger's causality in prices, thereby affecting the level of Google searches.

Thus, it is not significant evidence that Google searches affect prices, but that there is a greater chance that prices affect the curiosity of the general public. This shows that search engines are used after major price movements and not the other way around. The theory can be supported by the hypothesis that institutional investors, who are more representative, do not use common search engines - such as Google - as a source of information for decision making, corroborating the proposition that highly qualified mechanisms exist to assist them. them in your result.

Strong and moderate correlations were presented confirming how the maximum prices, like Petrobras, fluctuate and vary between search engines in Brazil and the world. These correlations confirm the idea that when prices reach their peaks - as well as their falls - they are validated in greater activation of the search curiosity for such companies on Google.

Even the behavior of different sectors, remaining uniform in this perspective, also presented moderate and positive correlations of strength, corroborating Granger's causality that prices affect search levels.

Another interesting fact is that another company that has a direct relationship with the population, but in another sector, Itaú (Financial Intermediaries), presented a weak and negative correlation, demonstrating again that we cannot affirm that there is a pattern in this sector.

Concerning the geographical filters between searches carried out by the World and searches carried out only in Brazil, it is not possible to state that there is any significant impact. The data alternates between companies, some of which are more or less related to global markets, but without a relevant difference. Thus, it is not possible to say that the national market is influenced by the levels of international searches, even though it is characterized by the large participation of foreigners. This corroborates the possibility that institutional investors do not use Google as a research source, only individual investors, who, in the world market and based on global data, are not very representative.

ACKNOWLEDGEMENTS

The authors are grateful to the UFABC Innovation and Sustainable Competitiveness Forum (FICS / UFABC) for intellectual support.

REFERENCES

[1] B3. Índice Bovespa (Ibovespa). Available in: http://www.b3.com.br/pt_br/market-data-e-indices/indices/indices-amplos/ibovespa.htm. Access: 20 Abr. 2020.

[2] BM&FBOVESPA. Dados de mercado. Available in: http://www.b3.com.br/pt_br/market-data-e-indices/servicos-de-dados/market-data/consultas/mercado-a-vista/dados-de-mercado Access: 20 Abr. 2020.

[3] Bordino, I., Battiston, S., Caldarelli, G., Cristelli, M., Ukkonen, A., & Weber, I. (2012). Web search queries can predict stock market volumes. PloS one, 7(7). https://doi.org/10.1371/journal.pone.0040014

[4] Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual web search engine.

[5] C.W.J. Granger Investigating causal relations by econometric models and cross-spectral methods Econometrica, 37 (1969), pp. 424-438 https://www.jstor.org/stable/1912791

[6] Challet, D., & Ayed, A. B. H. (2013). Predicting financial markets with Google Trends and not so random keywords. arXiv preprint arXiv:1307.4643.

[7] Curme, C., Preis, T., Stanley, H. E., & Moat, H. S. (2014). Quantifying the semantics of search behavior before stock market moves. Proceedings of the National Academy of Sciences, 111(32), 11600-11605. https://doi.org/10.1073/pnas.1324054111

[8] CVM. O que é a CVM?. CVM (2020). Available in: http://www.investidor.gov.br/menu/Menu_Investidor/a_cvm/ACVM.html. Access: 20 Abr. 2020.

[9] Da, Z., Engelberg, J., & Gao, P. (2011). In search of attention. The Journal of Finance, 66(5), 1461-1499. https://doi.org/10.1111/j.1540-6261.2011.01679.x

[10] Dimpfl, T., & Jank, S. (2016). Can internet search queries help to predict stock market volatility?. European Financial Management, 22(2), 171-192. https://doi.org/10.1111/eufm.12058

[11] Elder, A. (2004). Como Se Transformar Em Um Operador E Investidor... Gulf Professional Publishing.

[12] Ettredge, M., Gerdes, J., & Karuga, G. (2005). Using web-based search data to predict macroeconomic statistics. Communications of the ACM, 48(11), 87-92. https://dl.acm.org/doi/pdf/10.1145/1096000.1096010

[13] Facó, J. F. B., & Mandel, P. A. (2016). Aqui tem inovação!.Ed da Universidade Federal do ABC pag 53, São Bernardo do Campo.(ISBN: 978-85-68576-53-3)

[14] Faria, V. T.; Andrade, A. A.; Santos, J. P.; Facó, j. F. B.; Gasi, F. Measuring the impacts of database processing utilization in innovation processes on companies.

International journal of development research., v.10, p.34190 - 34194, 2020.(ISSN: 2230-9926)

[15] Galvão, A., Ribeiro É. "Mercado Financeiro: uma abordagem prática dos principais produtos e serviços." Elsevier Editora Ltda, Rio de Janeiro, 2006.(ISBN 978-85-352-1336-2

[16] Gil, A. C. (2019). Métodos e técnicas de pesquisa social. 7. ed. Ediitora Atlas SA.(ISBN 978-85-224-5142-5)

[17] Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S., & Brilliant, L. (2009). Detecting influenza epidemics using search engine query data. Nature, 457(7232), 1012-1014. https://doi.org/10.1038/nature07634

[18] GOOGLE. Trends Ajuda. Available in: http://support.google.com/trends. Access: 20 Abr. 2020.

[19] Gujarati, D. N., & Porter, D. C. (1999). Essentials of econometrics (Vol. 2). Singapore: Irwin/McGraw-Hill. (ISBN-13: 978-0-07-337584-7)

[20] IBOVESPA. Available in: http://www.b3.com.br/pt_br/market-data-e-indices/indices/indices-amplos/indice-ibovespa-ibovespa-composicao-da-carteira.htm. Access: 20 Abr. 2020.

[21] Infomoney 2020 - Petrobras (PETR3), Available in: https://www.infomoney.com.br/cotacoes/petrobras-petr3/ Access: 20 Abr. 2020.

[22] Internet Live Stats. Google Search Statistics. Available in: http://www.internetlivestats.com/google-search-statistics/#sources. Access: 20 Abr. 2020.

[23] Kleinberg, J. M. (1999). Authoritative sources in a hyperlinked environment. Journal of the ACM (JACM), 46(5), 604-632. https://doi.org/10.1145/324133.324140

[24] Lieberam-Schmidt, S. (2010). Analyzing and Influencing Search Engine Results: Business and Technology Impacts on Web Information Retrieval. Springer Science & Business Media.(ISBN:3834989150, 9783834989154)

[25] Limnios, C., & You, H. (2018). Can Google Trends actually improve housing market forecasts?. Available at SSRN 2886705. http://dx.doi.org/10.2139/ssrn.2886705

[26] Mao, H., Counts, S., & Bollen, J. (2011). Predicting financial markets: Comparing survey, news, twitter and search engine data. arXiv preprint arXiv:1112.1051.

[27] McKinney, W. (2012). Python for data analysis: Data wrangling with Pandas, NumPy, and IPython. " O'Reilly Media, Inc.".(ISBN: 978-1-449-31979-3)

[28] NET Statcounter GlobalStats - Marketshare. Search engine market share Brazil. Available in: http://gs.statcounter.com/search-engine-market-share/all/brazil#quarterly-200901-201702. Access: 20 Abr. 2020.

[29] Oust, A., & Martin, E. O. (2018). Can Google Search Data be Used as a Housing Bubble Indicator? (No. eres2018_151). European Real Estate Society (ERES).

[30] Popper, K. (2015). The Logic of Scientific Discovery. Central Works of Philosophy v4: Twentieth Century: Moore to Popper, 4, 262.(ISBN: 13: 978-1-84465-018-7)

[31] Preis, T., Moat, H. S., & Stanley, H. E. (2013). Quantifying trading behavior in financial markets using Google Trends. Scientific reports, 3, 1684. https://doi.org/10.1038/srep01684

[32] Preis, T., Reith, D., & Stanley, H. E. (2010). Complex dynamics of our economic life on different scales: insights from search engine query data. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 368(1933), 5707-5719. https://doi.org/10.1098/rsta.2010.0284

[33] Ramos, H. P., Ribeiro, K. K. M., & Perlin, M. S. (2017). The forecasting power of internet search queries in the Brazilian financial market. RAM. Revista de Administração Mackenzie, 18(2), 184-210. https://doi.org/10.1590/1678-69712017/administracao.v18n2p184-210

[34] Rao, T., & Srivastava, S. (2013, May). Modeling movements in oil, gold, forex and market indices using search volume index and twitter sentiments. In Proceedings of the 5th Annual ACM Web Science Conference (pp. 336-345). https://doi.org/10.1145/2464464.2464521

[35] Shim, S., Eastlick, M. A., Lotz, S. L., & Warrington, P. (2001). An online prepurchase intentions model: the role of intention to search: best overall paper award—The Sixth Triennial AMS/ACRA Retailing Conference, 2000☆. Journal of retailing, 77(3), 397-416. https://doi.org/10.1016/S0022-4359(01)00051-3

[36] Simon, H. A. (1955). A behavioral model of rational choice. The quarterly journal of economics, 69(1), 99-118. https://doi.org/10.2307/1884852

[37] Statcounter. Search engine market share Brazil. Available in: http://gs.statcounter.com/search-engine-market-share/all/brazil#quarterly-200901-201702. Access: 20 Abr. 2020.

[38] Sullivan, D. (2012). Google: 100 billion searches per month, search to integrate Gmail, launching enhanced search app for iOS. Search Engine Land.

[39] Wu, L., & Brynjolfsson, E. (2015). The future of prediction: How Google searches foreshadow housing prices and sales. In Economic analysis of the digital economy (pp. 89-118). University of Chicago Press. https://doi.org/10.7208/chicago/9780226206981.003.0003

[40] YAHOO, 2020. Bolsas e provedores de dados no Yahoo Finanças. Available in: https://br.ajuda.yahoo.com/kb/finance-for-web/SLN2310.html?impressions=true. Access: 20 Abr. 2020.

[41] Facó, J. F. B., de Andrade, A. A., & Gasi, F. (2020). Organizational Innovativeness in 21st Century: First Decade Construct Analysis. International Journal of Advanced Engineering Research and Science, 7(4). https://doi.org/10.22161/ijaers.74.6